

DAM
Desarrollo de Aplicaciones Multiplataforma
2º Curso

AD
Acceso a Datos

UD 7
Bases de Datos noSQL

IES BALMIS
Dpto Informática
Curso 2019-2020
Versión 1 (03/2019)

UD7 – Bases de Datos noSQL

ÍNDICE

1. Introducción
2. Ventajas de los sistemas noSQL
3. Diferencias con las Bases de Datos SQL
4. Tipos de Bases de Datos noSQL
5. Algunos ejemplos de Bases de Datos noSQL

1. Introducción

Hasta la fecha, la mayoría de los desarrolladores utilizaban sistemas de bases de datos SQL, pero de un tiempo a esta parte, un nuevo sistema de bases de datos parece haber llegado para quedarse, estos son los **sistemas de bases de datos NoSQL** (Not Only SQL – No sólo SQL).

Se puede decir que el término NoSQL aparece con la llegada de la Web 2.0. Hasta ese momento, sólo las grandes empresas subían contenidos a Internet si disponían de un portal web. Pero la aparición de plataformas como Facebook, Twitter, etc, cualquier usuario podía subir contenidos a Internet, lo que provocó un gran incremento de datos.

Es en este momento cuando comienzan a aparecer los primeros problemas en los sistemas de bases de datos relacionales. La primera opción fue ampliar máquinas, pero esta solución no funcionaba y resultaba muy costoso. La segunda opción fue crear un sistema específico a tal problema, obteniendo así una solución robusta al problema, de ahí surgen los sistemas de bases de datos NoSQL.

El uso de los sistemas de bases NoSQL permite superar los problemas de escalabilidad y rendimiento que se producen en los sistemas de bases de datos relacionales cuando se dan cita una gran cantidad de usuarios concurrentes y millones de consultas diarias. Estos sistemas de bases de datos NoSQL están optimizados para las tareas de recuperación y agregación de información.

Además, las bases de datos NoSQL son sistemas de almacenamiento de información que no cumplen el esquema Entidad-Relación, ni la estructura de tabla en la que se almacenan los datos. Estos sistemas almacenan la información haciendo uso de formatos como **Clave-Valor, mapeo de columnas o grafos**.

2. Ventajas de los sistemas NoSQL

La forma de almacenar de los sistemas de bases de datos NoSQL proporcionan una serie de ventajas sobre los sistemas de bases de datos relacionales.

- **Se pueden ejecutar sobre máquinas con pocos recursos:** estos sistemas, a diferencia de los sistemas basados en SQL, no requieren apenas computación, permitiendo así ejecutarse sobre máquinas de pocos recursos.
- **Escalabilidad horizontal:** para mejorar el rendimiento del sistema, únicamente es necesario añadir más nodos e indicar al sistema que nodos están disponibles.
- **Manejan grandes cantidades de datos:** esto es debido a que utilizan una estructura distribuida. En muchos casos se utilizan tablas Hash.
- **No generan cuellos de botella:** el principal problema de los sistemas SQL es que necesitan transcribir cada una de las sentencias para poder ejecutarlas. Esto

significa que hay un punto de entrada común y, ante muchas peticiones puede ralentizar el sistema.

- **Existen diferentes sistemas de bases de datos NoSQL** según la naturaleza de los proyectos, esto permite elegir la base de datos más óptima para cada necesidad.

3. Diferencias con las bases de datos SQL

Estas son algunas de las diferencias más destacables entre los sistemas de bases de datos NoSQL y los sistemas SQL.

- **No utilizan SQL como lenguaje de consultas:** la mayor parte de los sistemas de bases de datos NoSQL no utilizan este lenguaje o simplemente como apoyo. Por ejemplo, MongoDB utiliza JSON, BigTable utiliza GQL y Cassandra utiliza CQL.
- **No utilizan estructuras de almacenamiento fijas para guardar la información:** hacen uso de otros modelos para el almacenamiento de la información, como sistemas de Clave-Valor, grafos u objetos.
- **No suelen permitir operaciones JOIN:** al tener un volumen de datos extremadamente grande, resulta incómodo utilizar acciones con JOIN ya que la sobrecarga en el sistema puede resultar muy costosa. La solución a este problema puede estar en la desnormalización de los datos o realizar JOIN por software en la capa de aplicación.
- **Arquitectura distribuida:** las bases de datos relacionales suelen estar centralizadas en una sola máquina o con estructura Cliente-Servidor. Pero los sistemas de bases de datos NoSQL pueden contener la información compartida entre varias máquinas mediante tablas Hash distribuidas.

4. Tipos de bases de datos NoSQL

Según la forma que utilice de almacenar la información nos encontraremos con diferentes tipos de sistemas de bases de datos NoSQL, las más habituales son las que se muestran a continuación.

Clave-Valor

Estas son las habituales en los sistemas de bases de datos NoSQL ya que son las más sencillas en cuanto a funcionalidad se refiere. En este sistema, cada elemento es identificado por una clave única, permitiendo así una recuperación rápida. Además, la información es almacenada como un objeto binario (BLOB). Son muy eficientes en cuanto a lectura y escritura. Algunos ejemplos de este tipo son Cassandra, BigTable o Hbase.

Documentales

En el modelo documental la información se almacena como si de un documento se tratase, para ello utiliza una estructura simple como JSON o XML, en el cual se utilizará una clave única para cada registro. Además de realizar búsquedas por clave-valor, se pueden hacer búsquedas más avanzadas por el contenido del documento.

Este tipo de modelo es el más versátil, se puede utilizar en gran variedad de proyectos, incluyendo muchos que funcionarían con el modelo relacional. Algunos ejemplos serían MongoDB y CouchDB.

```
{
  Nombre:"Alberto",
  Dirección:"Castaños 17",
  Hijos:[
    {Nombre:"Andrés",    Edad:12},
    {Nombre:"Susana",    Edad:7},
    {Nombre:"Verónica",  Edad:4}
  ]
}
```

Grafos

En este modelo la información se representa como nodos de un grafos y sus relaciones como las aristas entre los nodos, de forma que para recorrer la información se puede aplicar la teoría de grafos. Para conseguir un rendimiento máximo la información debe estar normalizada.

Este modelo permite una navegación más eficiente entre relaciones que el modelo relacional. Algunos ejemplos son Neo4j, InfoGrid y Virtuoso.

Orientadas a objetos

Este modelo representa la información mediante objetos, al igual que ocurre en la programación orientada a objetos (POO) con JAVA, C# o Visual Basic .Net. Algunos ejemplos de este modelo son Zope, Gemstone, Matisse y Db4o.

Más información de tipos de Bases de Datos noSQL

<http://www.diegocalvo.es/base-de-datos-clave-valor/>

<http://www.diegocalvo.es/almacenes-en-columnas/>

<http://www.diegocalvo.es/base-de-datos-orientada-a-documentos/>

<http://www.diegocalvo.es/base-de-datos-orientada-a-grafos/>

5. Algunos ejemplos de bases de datos NoSQL

A continuación algunos ejemplos de sistemas de bases de datos NoSQL más utilizadas actualmente.

Cassandra



Base de datos creada por Apache de tipo Clave-Valor. Dispone de su propio lenguaje de consultas CQL (Cassandra Query Language). Es una aplicación JAVA y puede correr sobre cualquier plataforma que disponga de JVM.

Puedes visitar <http://cassandra.apache.org/> para más información.

MongoDB



Esta base de datos creada por 10gen Inc íntegramente en C++ proporciona gran velocidad a la hora de ejecutar sus operaciones. Es de esquema libre orientada a documento, lo que permite que entrada sea diferente del resto de registros almacenados.

MongoDB utiliza un lenguaje propio para el almacenamiento de la información, BSON que es una evolución de JSON, con la particularidad de poder almacenar datos binarios. Actualmente, MongoDB es la base de datos preferida entre los desarrolladores.

Puedes visitar <https://www.mongodb.com/es> para más información.

CouchDB



Creada por Apache y desarrollada utilizando el lenguaje de programación Erlang4 es capaz de funcionar sobre la mayoría de sistemas operativos. Destacar que utiliza como lenguaje de interacción JavaScript y JSON para el almacenamiento de archivos.

Permite la creación de vistas, lo que permite combinar documentos para obtener valores, es decir, permite realizar operaciones JOIN típicas de los sistemas de bases de datos relacionales.

Puedes visitar <http://couchdb.apache.org/> para más información.

Más ejemplos de BD noSQL

<http://nosql-database.org/>

Algunas empresas que utilizan este tipo de bases de datos

Estas son algunas de las grandes empresas de Internet que utilizan este tipo de bases de datos.

- **Cassandra**: Twitter, Facebook...
- **MongoDB**: FourSquare, SourceForge...
- **HBase**: Yahoo, Adobe...
- **Redis**: Flickr, Instagram, Github...
- **Neo4j**: Infojobs...

Cuando elegir sistemas de bases de datos NoSQL

- Cuando la escalabilidad en un modelo relacional no es viable, tanto a nivel técnico como de costes.
- Cuando el sistema tiene picos de uso muy elevados por parte de los usuarios.
- Cuando el número de datos crece rápidamente en situaciones puntuales, llegando a superar el Terabyte de información.
- Cuando la información no es homogénea, es decir, cuando en cada inserción podemos tener diferentes campos.